Casio Graphics Calculator
fx-9750GII
Guide for STAT 193

2013

This guide is for students in STAT193 using the fx-9750GII Casio Graphics Calculator.

Some students may have older versions of this calculator. Almost all of the functions of the fx-9750GII are the same, or nearly so, on older calculators. However there are some differences, and most importantly the older calculators cannot compute inverse T, inverse F and inverse Chi-square probabilities. Printed statistical tables must be used for these calculations.

A Supplementary Guide is available for students with older calculators explains what the differences are, and how to use statistical tables.

Contents

1 Turning the calculator on and off **3**

2 Important Keys **4**

3 Clearing the calculator's memory **4**

4 Mathematical Calculations **5**

5 Scientific Notation **6**

6 Entering, Editing and Deleting Data **7**

7 Summary Statistics for One Variable **8**

7.1 Individual Values 8

7.2 Grouped Data 9

8 Bivariate Data **10**

8.1 Entering bivariate data 10

8.2 Summary statistics for bivariate data, regression 10

8.3 Prediction using a fitted regression line 11

9 Probability Distributions **13**

9.1 Binomial Distribution 13

9.2 Normal Distribution 13

9.2.1 Normal Probabilities in Ranges 14

9.2.2 Inverse Normal Probabilities 15

9.3 t Distribution 17

9.3.1 t Probabilities in Ranges 17

9.3.2 Inverse t Probabilities 17

9.4 F Distribution 18

9.4.1 F Probabilities in Ranges 18

9.4.2 Inverse F Probabilities 18

9.5 Chi-square Distribution 19

9.5.1 Chi-square Probabilities in Ranges 19

9.5.2 Inverse Chi-square Probabilities 19

10 Large Sample Confidence Intervals **20**

10.1 Large Sample Confidence Interval for μ 20

10.2 Large Sample Confidence Interval for p 21

10.3 Large Sample Confidence Interval for $\mu_1 - \mu_2$ 21

11 Large Sample Hypothesis Tests **22**

11.1 Large Sample Hypothesis Test for μ 22

11.2 Large Sample Hypothesis Test for p 23

11.3 Large Sample Hypothesis Test for $\mu_1 - \mu_2$ 24

12 Small Sample Confidence Intervals **25**

13 Small Sample Hypothesis Tests **26**

14 Analysis of Variance – ANOVA **27**

15 Chi-Squared Test **29**

16 The Sign Test **30**

1 Turning the calculator on and off

- Turn on: AC/ON
- Turn off: SHIFT AC/ON

The calculator will turn off automatically after a short time without activity. No data are lost when the calculator does this.

2 Important Keys

- The [REPLAY] button – represented as ⊕ in these notes.

  Use it to move around on the screen when needed.

- The [MENU] key accesses the main functions of the calculator.

  Use ⊕ to move around to the panel you want, press [EXE] to select.

  We will only need:

  - RUN-MAT: For mathematical calculations;
  - STAT: For statistical calculations;
  - SYSTEM: To clear the calculator's memory.

- At any stage the [EXIT] key will take you back one step, keep pressing it and you'll end up at the beginning screen for whichever [MENU] mode you're in.

- You can switch backwards and forwards between RUN-MAT and STAT modes, and won't lose any data.

  Do this when you want to do a mathematical calculation (using RUN-MAT mode) midway through a statistical analysis (using STAT mode).

3 Clearing the calculator's memory

You will need to do this at the beginning of the final examination and in the term test.

It's also worth clearing the menu every time you start a new statistical problem – it will clear out any data that might be lurking in the calculator's memory.

- [MENU] then move (using ⊕) to the SYSTEM panel; [EXE]
- [F5] Reset
- [F2] Main Memories
- [F1] Yes – to rest the main memories
  (this can take a few seconds)
- [EXIT]
- Then [MENU] to get you back to the menus to start working again.

NB: If you need to change settings on your calculator for some calculations, clearing the memory is a good way of making sure that everything is back to the way it was before you made any changes, and that all the data are gone.

4 Mathematical Calculations

Use your calculator in RUN-MAT mode: [MENU]; ⊕ RUN-MAT; [EXE].

Type in mathematical expressions, and evaluate using [EXE].

The answer appears at right.

Examples:

$2+2$	[2] [+] [2] [EXE]
4.2^2	[4] [.] [2] [x^2] [EXE]
$\frac{6.2-3.1}{3.4}$	[(] [6] [.] [2] [−] [3] [.] [1] [)] [÷] [3] [.] [4] [EXE]

Be careful about when you need to use brackets [(] and [)]. Without brackets the last example would evaluate $6.2-\frac{3.1}{3.4}$.

You can re-use the answer from one calculation in the next using [SHIFT] [Ans].

Another example showing how to do a calculation in two stages using [SHIFT] [Ans]:

$\frac{8-3.2}{\sqrt{4+2}}$	[(] [8] [−] [3] [.] [2] [)] [÷] [SHIFT] [$\sqrt{\ }$] [(] [4] [+] [2] [)] [EXE]
OR	
$\frac{8-3.2}{\sqrt{4+2}}$	[8] [−] [3] [.] [2] [EXE] [SHIFT] [Ans] [÷] [SHIFT] [$\sqrt{\ }$] [(] [4] [+] [2] [)] [EXE]

NB: You can move in and out of RUN-MAT mode even when you're in the middle of another calculation in STAT mode. Use the [MENU] key to switch between them.

5 Scientific Notation

There is a standard way of expressing numbers that may be very large or very small.

Powers of ten are:

$$10^0 = 1$$

$10^1 = 10$	$10^{-1} = 0.1$
$10^2 = 100$	$10^{-2} = 0.01$
$10^3 = 1000$	$10^{-3} = 0.001$
$10^4 = 10000$	$10^{-4} = 0.0001$
$10^5 = 100000$	$10^{-5} = 0.00001$
$10^6 = 1000000$	$10^{-6} = 0.000001$
$\vdots$	$\vdots$

Use these powers of 10 to represent numbers in a compact form:

Number		Scientific Notation		Calculator
10334.5	$= 1.03345 \times 10000$	$= 1.03345 \times 10^4$		`1.03345E+04`
0.000024	$= 2.4 \times 0.00001$	$= 2.4 \times 10^{-5}$		`2.4E-05`
1	$= 1 \times 1$	$= 1 \times 10^0$		`1.0E+00`

The EXP button on a calculator makes the 'E' when entering numbers in scientific notation.

Type		To get...
`1E5`	1×10^5	100000
`2.48E6`	2.48×10^6	2480000
`6.3E-1`	6.3×10^{-1}	0.63
`4.443E-3`	4.443×10^{-3}	0.00443

6 Entering, Editing and Deleting Data

Use your calculator in STAT mode: [MENU]; ⊕ STAT; [EXE].

You'll see the List layout:

	List 1	List 2	List 3	List 4
SUB				
1				
2				
3				
⋮	⋮			

You can move around this from one cell to another using ⊕.

There is a set of menu items above the six function keys [F1], [F2], . . . , [F6]:

GRPH CALC TEST INTR DIST ▷

[F1] [F2] [F3] [F4] [F5] [F6]

To select the menu item, press the function key directly below it.

For example [F2] gives the CALC function. In these notes this is represented by [F2] CALC.

Pressing [F6] ▷ gives a second set of function options

TOOL EDIT DEL DEL-A INS ▷

[F1] [F2] [F3] [F4] [F5] [F6]

And pressing [F6] ▷ again gives the original set.

- To enter data, move to the cell you want, type the number you want to enter, and press [EXE];
  The cursor will move down to the next cell below.
- To delete an item move to the cell you want to delete, and press [DEL].
- To clear an entire list move the cursor into the list you want to clear, and press [F4] DEL-A (you may have to press [F6] ▷ first to find DEL-A).
  Press [F1] Yes to confirm that you do want to clear the data.

7 Summary Statistics for One Variable

Use your calculator in STAT mode: [MENU]; ⊕ STAT; [EXE].

7.1 Individual Values

Data are the $n = 9$ values: 2, 5, 5, 8, 8, 8, 11, 11, 14

Input x_i's individually into List 1: 2 [EXE] 5 [EXE] 5 [EXE] ...

	List 1	List 2	List 3	List 4
SUB				
1	2			
2	5			
3	5			
⋮	⋮			

- [F2] CALC
- [F1] 1VAR

Read off the sample statistics:

1-Variable		
$\bar{x}$	=8	mean
$\sum x$	=72	sum $\sum_i x_i$
$\sum x^2$	=684	sum $\sum_i x_i^2$
σx	=3.46410161	Std. Dev. σ (n)
sx	=3.67423461	Std. Dev. s ($n-1$) ← use this one
n	=9	Sample size
minX	=2	Smallest value
Q1	=5	Lower Quartile (LQ)
Med	=8	Median
Q3	=11	Upper Quartlie (UQ)
maxX	=14	Largest value
Mod	=8	Mode
Mod:n	=1	Number of modes
Mod:F	=3	Frequency at the mode

No matter which statistic you want, always check the sample size is right!

7.2 Grouped Data

Input x_i's using frequencies:

x_i	2	5	8	11	14
f_i	1	2	3	2	1

Enter the x_i's into List 1: 2 [EXE] 5 [EXE] 8 [EXE] 11 [EXE] 14
Enter the f_i's into List 2: 1 [EXE] 2 [EXE] 3 [EXE] 2 [EXE] 1

	List 1	List 2	List 3	List 4
SUB				
1	2	1		
2	5	2		
3	8	3		
4	11	2		
5	14	1		

The calculator assumes that List 1 is x, but we have to tell it that List 2 is f

- [F2] CALC
- [F6] SET
- $\bigoplus$: 1Var XList; [F1] LIST; [1]; [EXE]
- $\bigoplus$: 1Var Freq; [F2] LIST; [2]; [EXE]
- [EXIT]
- [F1] 1VAR

Always check the sample size is right! $n = 9$

You may need to set these list assignments back: best to clear the memories if you are unsure of any setting.

8 Bivariate Data

8.1 Entering bivariate data

Example. Diameters at breast height of 14 chestnut oak trees for a sample of such trees grown in poor soil.

Age (years)	DBH (inches)	Age (years)	DBH (inches)
x_i	y_i	x_i	y_i
4	0.8	23	4.7
8	1.0	28	6.0
8	3.0	30	6.0
10	3.5	33	8.0
13	3.5	35	7.0
16	4.5	38	7.0
20	5.5	42	7.5

Enter the data into Lists 1 and 2 in your calculator (STAT mode).

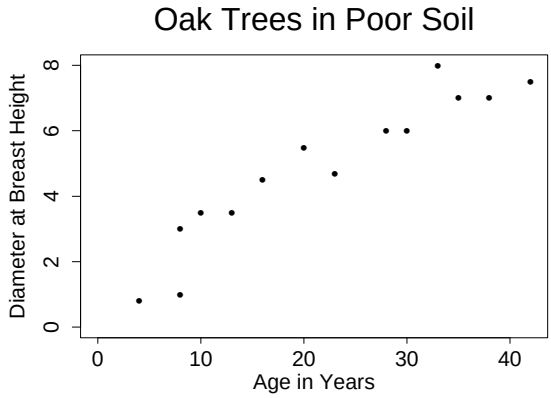
	List 1	List 2	List 3	List 4
SUB				
1	4	0.8		
2	8	1.0		
3	8	3.0		
⋮	⋮			
14	42	7.5		

8.2 Summary statistics for bivariate data, regression

Once the data have been entered:

Draw a scatterplot of the data:

- F1 GRPH
- F2 GPH2



Then fit the line:

- [EXIT] [EXIT] back to the data
- [F2] CALC
- [F3] REG
- [F1] X
- [F2] $a+bX \leftarrow$ **NB**: use this option or a, b are reversed

Reports statistics from linear regression:

a	=1.0841	Intercept
b	=0.1715	Slope
r	=0.9387	Pearson's Correlation Coefficient
r^2	=0.8811	Coefficient of Determination
MSE	=0.6706	s^2

NOTE: To calculate Spearman's correlation coefficient, replace the data with their ranks (taking care to treat ties properly), and compute Pearson's correlation coefficient for the ranks.

8.3 Prediction using a fitted regression line

After fitting the regression line go into RUN-MAT mode and set the calculator up to make predictions:

- [MENU]; $\oplus$ RUN-MAT; [EXE].
- [OPTN]
- [F5] STAT

Then to predict Y for a particular value of X, e.g. $X = 8$:

- Type X value: [8]
- [F2] $\widehat{Y}$
- [EXE]

Fitted value: $\widehat{Y} = 2.4562$.

$(8, 2.4562)$ is on the regression line.

Alternative Method for Prediction

If the data are already entered then graph the regression line

- F1 GRPH
- F2 GPH2
- F1 CALC
- F2 X
- F2 $a+bX \leftarrow$ **NB**: use this option or a, b are reversed
- F6 DRAW

Now predict the points – go into prediction mode 'G-Solv':

- SHIFT F5 G-Solv
- F1 Y-CAL

Can now predict Y for any X: e.g. $X = 8$:

- Type X value: 8 EXE

Fitted value: $\widehat{Y} = 2.4562$.

$(8, 2.4562)$ is on the regression line.

9 Probability Distributions

9.1 Binomial Distribution

If $X \sim \text{Binomial}(n, p)$ then the calculator can compute

- Individual Probabilities: $P(X = x)$
- Cumulative Probabilities: $P(X \leq x)$

Example.

$$X \sim \text{Binomial}(15, 0.05)$$

Individual probability: compute $P(X = 3)$:

- [F5] DIST
- [F5] BINM
- [F1] Bpd
  ('p' for probability)
- ⊕ Data; [F2] Var
- ⊕ x; [3] [EXE]
- ⊕ Numtrial; [1] [5] [EXE]
- ⊕ p; [0] [.] [0] [5] [EXE]
- ⊕ Execute; [EXE]

Answer is 0.0307

Cumulative probability: compute $P(X \leq 2)$:

- [F5] DIST
- [F5] BINM
- [F2] Bcd
  ('c' for cumulative)
- ⊕ Data; [F2] Var
- ⊕ x; [2] [EXE]
- ⊕ Numtrial; [1] [5] [EXE]
- ⊕ p; [0] [.] [0] [5] [EXE]
- ⊕ Execute; [EXE]

Answer is 0.9638

To calculate $P(X \geq x)$ use

$$\begin{aligned} P(X \geq x) &= 1 - P(X < x) \\ &= 1 - P(X \leq (x-1)) \end{aligned}$$

e.g. $P(X \geq 3) = 1 - P(X \leq 2) = 1 - 0.9638 = 0.0362$.

Note: Can switch to RUN-MAT mode:
[1] [-] [SHIFT] [Ans] [EXE]
[Ans] uses the last result from STAT mode 0.9638.

9.2 Normal Distribution

If $X \sim \text{Normal}(\mu, \sigma^2)$ then the calculator can compute

- Probabilities in ranges: $P(a < X < b)$
- Inverse probabilities: given a probability p find:
  - (Left): x^* so that $P(X < x^*) = P(-\infty < X < x^*) = p$
  - (Right): x^* so that $P(X > x^*) = P(x^* < X < \infty) = p$
  - (Central): (x_1^*, x_2^*) so that $P(x_1^* < X < x_2^*) = p$

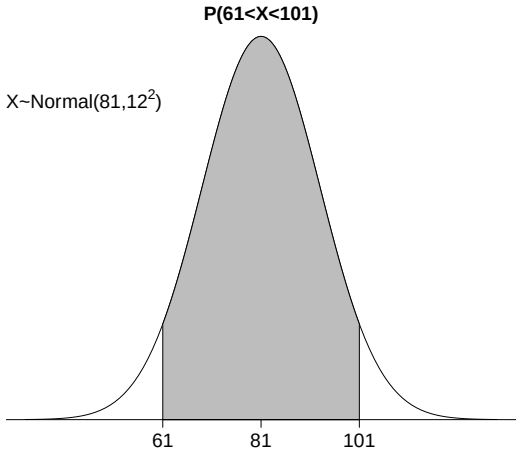
9.2.1 Normal Probabilities in Ranges

NB: Use $\pm 9 \times 10^{99} = \pm$ `9E99` for plus or minus infinity ($\pm\infty$) when calculating left and right tail probabilities.

Example: Lengths of oysters follow a normal distribution with mean 81mm and standard deviation 12mm.

$$X \sim \text{Normal}(81, 12^2)$$

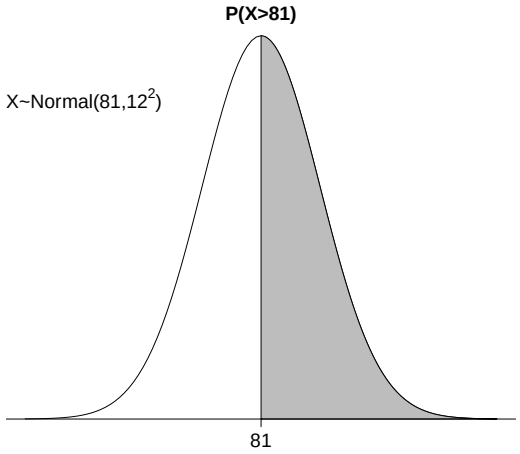
Find $P(61 < X < 101)$



- [F5] DIST
- [F1] NORM
- [F2] Ncd
  ('c' for cumulative)
- $\oplus$ Data; [F2] Var
- $\oplus$ Lower; [6] [1] [EXE]
- $\oplus$ Upper; [1] [0] [1] [EXE]
- $\oplus$ σ; [1] [2] [EXE]
- $\oplus$ μ; [8] [1] [EXE]
- $\oplus$ Execute; [EXE]

Answer is 0.90442

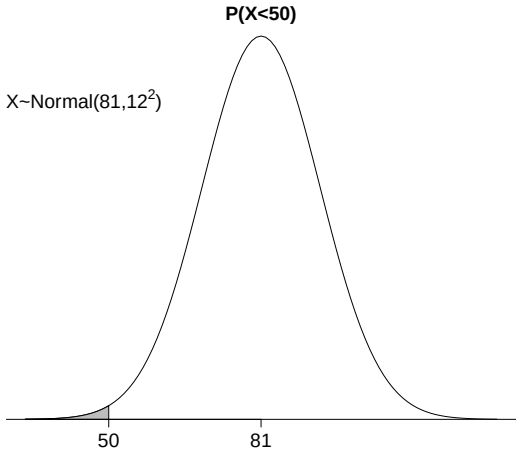
Find $P(X > 81)$



- [F5] DIST
- [F1] NORM
- [F2] Ncd
  ('c' for cumulative)
- $\oplus$ Data; [F2] Var
- $\oplus$ Lower; [8] [1] [EXE]
- $\oplus$ Upper; [9] [EXP] [9] [9] [EXE]
- $\oplus$ σ; [1] [2] [EXE]
- $\oplus$ μ; [8] [1] [EXE]
- $\oplus$ Execute; [EXE]

Answer is 0.50000

Find $P(X < 50)$



- F5 DIST
- F1 NORM
- F2 Ncd
  ('c' for cumulative)
- ⊕ Data; F2 Var
- ⊕ Lower; – 9 EXP
  9 9 EXE
- ⊕ Upper; 5 0 EXE
- ⊕ σ; 1 2 EXE
- ⊕ μ; 8 1 EXE
- ⊕ Execute; EXE

Answer is $0.00489 = 4.8925 \times 10^{-3} = 4.8925\text{E}-03$

Note: When calculating $P(x_1 < X < x_2)$ for $X \sim \text{Normal}(\mu, \sigma^2)$, the calculator also reports the standardised versions (z_1, z_2) of (x_1, x_2):

$$z_1 = \frac{x_1 - \mu}{\sigma} \quad \text{and} \quad z_2 = \frac{x_2 - \mu}{\sigma}$$

for which

$$P = P(x_1 < X < x_2) = P(z_1 < Z < z_2)$$

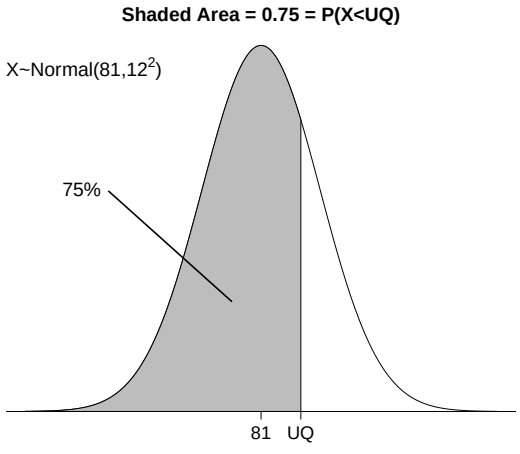
with $Z \sim \text{Normal}(0, 1)$.

9.2.2 Inverse Normal Probabilities

Example: Lengths of oysters follow a normal distribution with mean 81mm and standard deviation 12mm.

$$X \sim \text{Normal}(81, 12^2)$$

Find x^* so that $P(X < x^*) = 0.75$
i.e. the Upper Quartile of the distribution of X.
(The area 0.75 is to the Left)

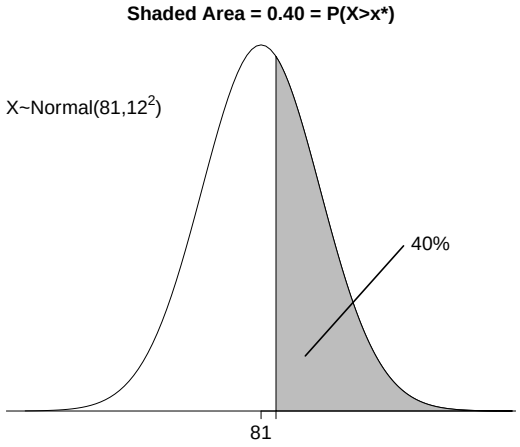


- F5 DIST
- F1 NORM
- F3 InvN
  (Inverse Normal)
- ⊕ Data; F2 Var
- ⊕ Tail; F1 Left
- ⊕ Area; 0 . 7 5 EXE
- ⊕ σ; 1 2 EXE
- ⊕ μ; 8 1 EXE
- ⊕ Execute; EXE

Answer is $x^* = 89.09$

Find x^* so that $P(X > x^*) = 0.40$
(The area 0.40 is to the Right)

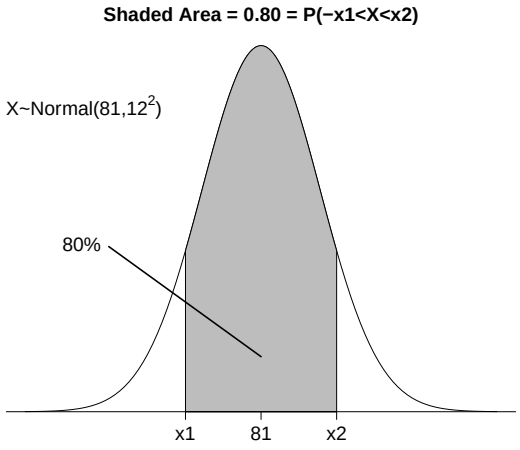
- [F5] DIST
- [F1] NORM
- [F3] InvN
  (Inverse Normal)
- $\oplus$ Data; [F2] Var
- $\oplus$ Tail; [F2] Right
- $\oplus$ Area; [0] [.] [4] [0] [EXE]
- $\oplus$ σ; [1] [2] [EXE]
- $\oplus$ μ; [8] [1] [EXE]
- $\oplus$ Execute; [EXE]



Answer is $x^* = 84.04$

Find (x_1^*, x_2^*) so that $P(x_1^* < X < x_2^*) = 0.80$
(The area 0.80 is at the Centre)

- [F5] DIST
- [F1] NORM
- [F3] InvN
  (Inverse Normal)
- $\oplus$ Data; [F2] Var
- $\oplus$ Tail; [F3] CNTR
- $\oplus$ Area; [0] [.] [8] [0] [EXE]
- $\oplus$ σ; [1] [2] [EXE]
- $\oplus$ μ; [8] [1] [EXE]
- $\oplus$ Execute; [EXE]



Answer is $(x_1^*, x_2^*) = (65.621, 96.379)$

9.3 t Distribution

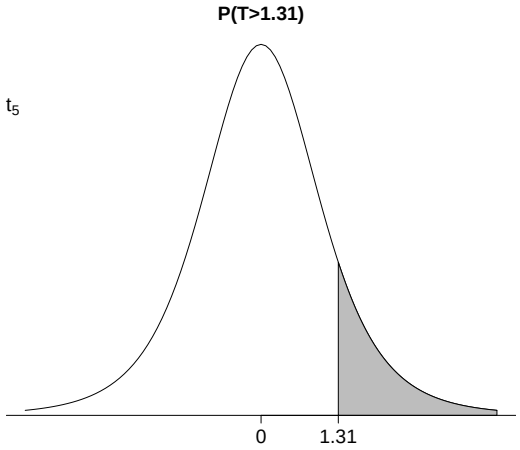
If $T \sim t_d$ = a t-distribution with d degrees of freedom then the calculator can compute

- Probabilities in ranges: $P(a < T < b)$
- Inverse probabilities: given a probability p find:
  - (Right): t_d^* so that $P(T > t_d^*) = P(t_d^* < X < \infty) = p$

9.3.1 t Probabilities in Ranges

NB: Use $\pm 9 \times 10^{99} = \pm$`9E99` for plus or minus infinity ($\pm\infty$) when calculating left and right tail probabilities.

Find $P(t_{12} > 1.31)$

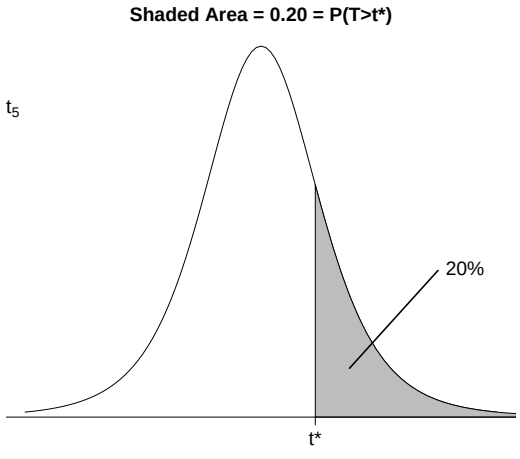


- F5 DIST
- F2 t
- F2 tcd
  ('c' for cumulative)
- ⊕ Data; F2 Var
- ⊕ Lower; 1 . 3 1 EXE
- ⊕ Upper; 9 E 9 9 EXE
- ⊕ df; 1 2 EXE
- ⊕ Execute; EXE

Answer is 0.10736

9.3.2 Inverse t Probabilities

Find t^* so that $P(t_{12} > t^*) = 0.20$



- F5 DIST
- F2 t
- F3 Invt
- ⊕ Data; F2 Var
- ⊕ Area; 0 . 2 0 EXE
- ⊕ Upper; 9 E 9 9 EXE
- ⊕ df; 1 2 EXE
- ⊕ Execute; EXE

Answer is $t^* = 0.87261$

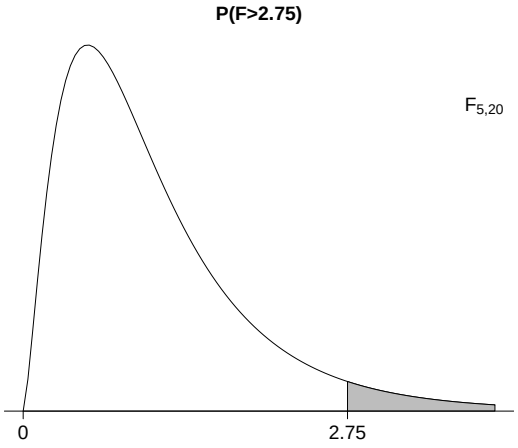
9.4 F Distribution

If $F \sim F_{(d1,d2)}$ = an F-distribution with (d_1, d_2) degrees of freedom then the calculator can compute

- Probabilities in ranges: $P(a < F < b)$
- Inverse probabilities: given a probability p find:
  - (Right): $F^*_{(d1,d2)}$ so that $P(F > F^*_{(d1,d2)}) = P(F^*_{(d1,d2)} < F < \infty) = p$

9.4.1 F Probabilities in Ranges

Find $P(F_{(5,12)} > 2.75)$

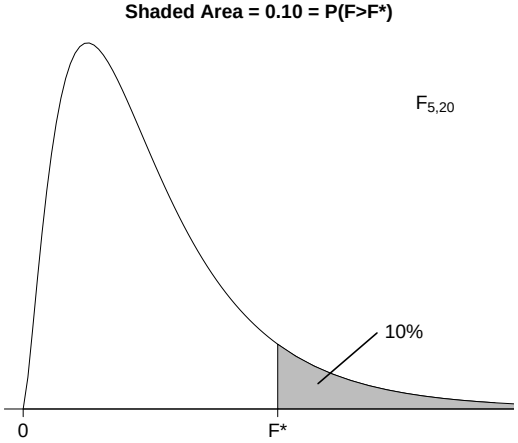


- F5 DIST
- F4 F
- F2 Fcd
  ('c' for cumulative)
- ⊕ Data; F2 Var
- ⊕ Lower; 2 . 7 5 EXE
- ⊕ Upper; 9 E 9 9 EXE
- ⊕ n:df; 5 EXE (numerator df=d1)
- ⊕ d:df; 1 2 EXE (denominator df=d2)
- ⊕ Execute; EXE

Answer is 0.07014

9.4.2 Inverse F Probabilities

Find F^* so that $P(F_{(5,12)} > F^*) = 0.10$



- F5 DIST
- F4 F
- F3 InvF
- ⊕ Data; F2 Var
- ⊕ Area; 0 . 1 EXE;
- ⊕ n:df; 5 EXE (numerator df=d1)
- ⊕ d:df; 1 2 EXE (denominator df=d2)
- ⊕ Execute; EXE

Answer is $F^* = 2.3940$

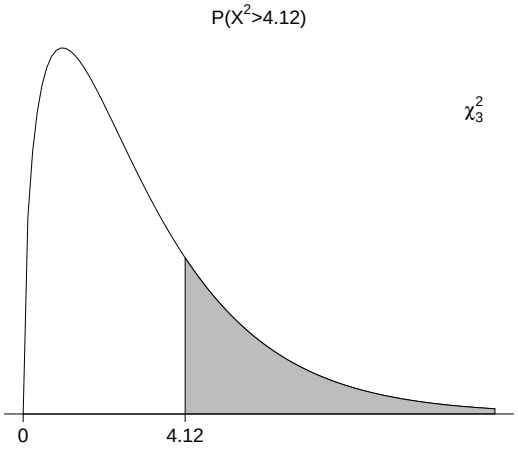
9.5 Chi-square Distribution

If $X^2 \sim \chi^2_d$ = a χ^2-distribution with d degrees of freedom then the calculator can compute

- Probabilities in ranges: $P(a < X^2 < b)$
- Inverse probabilities: given a probability p find:
  - (Right): χ^{2*}_d so that $P(X^2 > \chi^{2*}_d) = P(\chi^{2*}_d < X^2 < \infty) = p$

9.5.1 Chi-square Probabilities in Ranges

Find $P(\chi^2_3 > 4.12)$

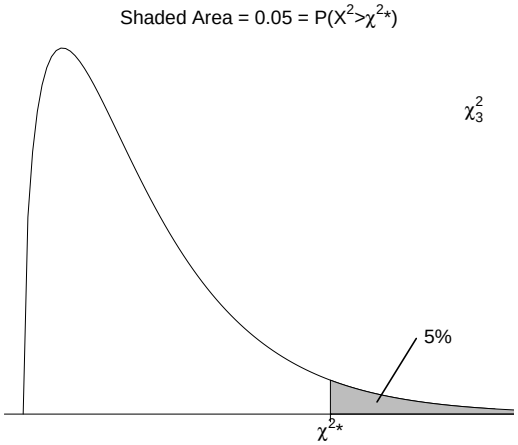


- F5 DIST
- F3 CHI
- F2 Ccd
  ('c' for cumulative)
- ⊕ Data; F2 Var
- ⊕ Lower; 4 . 1 2 EXE
- ⊕ Upper; 9 E 9 9 EXE
- ⊕ df; 3 EXE
- ⊕ Execute; EXE

Answer is 0.24879

9.5.2 Inverse Chi-square Probabilities

Find χ^{2*} so that $P(X^2_3 > \chi^{2*}) = 0.05$



- F5 DIST
- F3 CHI
- F3 InvC
- ⊕ Data; F2 Var
- ⊕ Area; 0 . 0 5 EXE
- ⊕ df; 3 EXE
- ⊕ Execute; EXE

Answer is 7.8147

10 Large Sample Confidence Intervals

Large sample confidence intervals can be calculated for means, differences in means, proportions, differences in proportions when sample sizes are **large** $n \geq 30$. For two sample intervals (differences in means or proportions) **both** samples must be large ($n_1 \geq 30$ and $n_2 \geq 30$).

To compute a large sample confidence interval

- [F4] INTR (Interval)
- [F1] Z (Large sample)
- Choose one of:
  - [F1] 1-S: CI for μ (mean of one pop.)
  - [F2] 2-S: CI for $\mu_1 - \mu_2$ (mean difference between two pops.)
  - [F3] 1-P: CI for p (proportion in one pop.)
  - [F4] 2-P: CI for $p_1 - p_2$ (difference in proportion between two pops.)
- ... Specify other information depending on the choice above
  including the confidence level (e.g. 95%)

The calculator produces a pair of numbers: (Left, Right) = the confidence interval.

10.1 Large Sample Confidence Interval for μ

Example. When $\bar{x} = 27.6$, $s = 4$ and $n = 40$.

Create a 95% confidence interval for the population mean μ.

- [F4] INTR (Interval)
- [F1] Z (Large sample)
- [F1] 1-S (One sample)
- $\bigoplus$ Data; [F2] Var
- $\bigoplus$ C-Level; [0] [.] [9] [5] [EXE] (Confidence level)
- $\bigoplus$ σ; [4] [EXE] (This is s)
- $\bigoplus$ $\bar{x}$; [2] [7] [.] [6] [EXE]
- $\bigoplus$ n; [4] [0] [EXE]
- $\bigoplus$ Execute; [EXE]

Gives $(26.36, 28.84)$.

10.2 Large Sample Confidence Interval for p

In addition to having $n \geq 30$ must also check that

$$p \pm 3\sqrt{\frac{p(1-p)}{n}}$$

is inside the range (0,1). A quick way to check this is to construct a 99.7% interval, and see if it is in the range (0,1). (The calculator is dumb enough to create confidence intervals outside this range without reporting an error!)

Example. When $X = 26$ and $n = 40$.

Create a 98% confidence interval for the population proportion p.

- F4 INTR (Interval)
- F1 Z (Large sample)
- F3 1-P (One proportion)
- $\bigoplus$ C-Level; 0 . 9 8 EXE (Confidence level)
- $\bigoplus$ x; 2 6 EXE
- $\bigoplus$ n; 4 0 EXE
- $\bigoplus$ Execute; EXE

Gives $(0.475, 0.825)$.

10.3 Large Sample Confidence Interval for $\mu_1 - \mu_2$

Example.

$$n_1 = 50 \qquad \bar{x}_1 = 26.1 \qquad s_1 = 3.6$$
$$n_2 = 40 \qquad \bar{x}_2 = 30.5 \qquad s_2 = 5.4$$

Create a 95% confidence interval for $\mu_1 - \mu_2$

- F4 INTR (Interval)
- F1 Z (Large sample)
- F2 2-S (Two samples)
- $\bigoplus$ Data; F2 Var
- $\bigoplus$ C-Level; 0 . 9 5 EXE (Confidence level)
- $\bigoplus$ $\sigma 1$; 3 . 6 EXE (This is s_1)
- $\bigoplus$ $\sigma 2$; 5 . 4 EXE (This is s_2)
- $\bigoplus$ $\bar{x}1$; 2 6 . 1 EXE
- $\bigoplus$ $n1$; 5 0 EXE
- $\bigoplus$ $\bar{x}2$; 3 0 . 5 EXE
- $\bigoplus$ $n2$; 4 0 EXE
- $\bigoplus$ Execute; EXE Gives $(-6.35, -2.45)$

11 Large Sample Hypothesis Tests

Large sample hypothesis tests can be carried out for means, differences in means, proportions, differences in proportions when sample sizes are **large** $n \geq 30$. In two sample tests **both** samples must be large ($n_1 \geq 30$ and $n_2 \geq 30$).

To carry out a large sample hypothesis test

- F3 TEST
- F1 Z (Large sample)
- Choose one of:
  - F1 1-S: Test $H_0 : \mu = \mu_0$ (mean of one pop.)
  - F2 2-S: Test $H_0 : \mu_1 - \mu_2 = 0$ (mean difference between two pops.)
  - F3 1-P: Test $H_0 : p = p_0$ (proportion in one pop.)
  - F4 2-P: Test $H_0 : p_1 - p_2 = 0$ (difference in proportion between two pops.)
- ... Specify other information depending on the choice above
  including the type of alternative hypothesis (one or two sided)

The calculator produces a test statistic value (Z) and a p-value, as well as other information depending on the chosen test.

Use the p-value to determine the outcome of the test at chosen significance level α (reject H_0 if p-value $< \alpha$).

To determine the **critical value** Z^* for the test statistic Z which specifies the rejection region, use the Inverse Normal function of the calculator.

The rejection region may be one or two sided depending on H_1. Find Z^* so that

Two sided	$P(Z > Z^*) = \frac{\alpha}{2}$
One sided (right)	$P(Z > Z^*) = \alpha$
One sided (left)	$P(Z < Z^*) = \alpha$

(Always sketch the rejection region)

To find Z^* use the methods in Section 9.2.2.

11.1 Large Sample Hypothesis Test for μ

Example. Have sample information $\bar{x} = 10.2$, $s = 2.3$ and $n = 30$.

Test the hypotheses

$$H_0 : \mu = 9.4 \qquad H_1 : \mu \neq 9.4$$

A two-sided test.

- F3 TEST
- F1 Z (Large sample)
- F1 1-S (One sample)
- $\bigoplus$ Data; F2 Var
- $\bigoplus$ μ; F1 $\neq \mu 0$ (2 sided)

- $\bigoplus$ $\mu 0$; [9] [.] [4] [EXE]
- $\bigoplus$ σ; [2] [.] [3] [EXE]
- $\bigoplus$ $\bar{x}$; [1] [0] [.] [2] [EXE]
- $\bigoplus$ n; [3] [0] [EXE]
- $\bigoplus$ Execute; [EXE]

Result: Test Statistic, $Z = 1.905$, p-value $= 0.05676$.

Critical value Z^* for a two-sided $\alpha = 5\%$ test: find Z^* so that $P(Z > Z^*) = 0.05/2 = 0.025$ for $Z \sim N(0,1)$:

- [F5] DIST
- [F1] NORM
- [F3] InvN
  (Inverse Normal)
- $\bigoplus$ Data; [F2] Var
- $\bigoplus$ Tail; [F2] Right
- $\bigoplus$ Area; [0] [.] [0] [2] [5] [EXE]
- $\bigoplus$ σ; [1] [EXE]
- $\bigoplus$ μ; [0] [EXE]
- $\bigoplus$ Execute; [EXE]

Answer is $Z^* = 1.96$.

The rejection region is $Z < -1.96$ or $Z > +1.96$.

11.2 Large Sample Hypothesis Test for p

In addition to having $n \geq 30$ must also check that

$$p \pm 3\sqrt{\frac{p(1-p)}{n}}$$

is inside the range (0,1). A quick way to check this is to construct a 99.7% confidence interval for p, and see if it is in the range (0,1). (The calculator is dumb enough to create confidence intervals outside this range without reporting an error!)

Example. Have sample information $x = 26$ successes in $n = 40$ independent Binomial trials.

Test the hypotheses

$$H_0 : p = 0.60 \qquad H_1 : p > 0.60$$

A one-sided test.

- [F3] TEST
- [F1] Z (Large sample)
- [F3] 1-P (One proportion)
- $\bigoplus$ Prop; [F3] > (1 sided, >)

- $\bigoplus p0$; [0] [.] [6] [EXE]
- $\bigoplus \bar{x}$; [2] [6] [EXE]
- $\bigoplus n$; [4] [0] [EXE]
- $\bigoplus$ Execute; [EXE]

Result: Test Statistic, $Z = 0.6455$, p-value $= 0.25930$.

11.3 Large Sample Hypothesis Test for $\mu_1 - \mu_2$

Example. Given data

$$\begin{array}{lll} n_1 = 50 & \bar{x}_1 = 26.1 & s_1 = 3.6 \\ n_2 = 40 & \bar{x}_2 = 30.5 & s_2 = 5.4 \end{array}$$

Test the hypotheses:

$$H_0 : \mu_1 = \mu_2 \qquad H_1 : \mu_1 < \mu_2$$

One sided.

- [F3] TEST
- [F1] Z (Large sample)
- [F2] 2-S (Two Samples)
- $\bigoplus$ Data; [F2] Var
- $\bigoplus \mu 1$; [F2] $< \mu 2$ (1 sided, $<$)
- $\bigoplus \sigma 1$; [3] [.] [6] [EXE] (This is s_1)
- $\bigoplus \sigma 2$; [5] [.] [4] [EXE] (This is s_2)
- $\bigoplus \bar{x}1$; [2] [6] [.] [1] [EXE]
- $\bigoplus n1$; [5] [0] [EXE]
- $\bigoplus \bar{x}2$; [3] [0] [.] [5] [EXE]
- $\bigoplus n2$; [4] [0] [EXE]
- $\bigoplus$ Execute; [EXE]

Result: Test Statistic, $Z = -4.426$, p-value $= \texttt{4.7956E} - \texttt{06} = 0.00000479$.

12 Small Sample Confidence Intervals

For small samples ($n < 30$) of **Normally distributed** data.

To compute a small sample confidence interval

- [F4] INTR (Interval)
- [F2] t (Small sample)
- Choose one of:
  - [F1] 1-S: CI for μ (mean of one pop.)
  - [F2] 2-S: CI for $\mu_1 - \mu_2$ (mean difference between two pops.)
- ... Specify other information depending on the choice above **including** the confidence level (e.g. 95%)

The calculator produces a pair of numbers: (Left, Right) = the confidence interval.

Example. When $\bar{x} = 30.6$, $sx = 9.5$ and $n = 20$.

Create a 95% confidence interval for the population mean μ.

- [F4] INTR (Interval)
- [F2] t (Small sample)
- [F1] 1-S (One sample)
- $\bigoplus$ Data; [F2] Var
- $\bigoplus$ C-Level; [0] [.] [9] [5] [EXE] (Confidence level)
- $\bigoplus$ $\bar{x}$; [3] [0] [.] [6] [EXE]
- $\bigoplus$ sx; [9] [.] [5] [EXE]
- $\bigoplus$ n; [2] [0] [EXE]
- $\bigoplus$ Execute; [EXE]

Gives $(26.2, 35.0)$.

13 Small Sample Hypothesis Tests

For small samples ($n < 30$) of **Normally distributed** data.

To carry out a small sample hypothesis test

- F3 TEST
- F2 t (small sample)
- Choose one of:
  - F1 1-S: Test $H_0 : \mu = \mu_0$ (mean of one pop.)
  - F2 2-S: Test $H_0 : \mu_1 - \mu_2 = 0$ (mean difference between two pops.)
  - F3 REG: Test $H_0 : \beta = 0$ (slope of regression line)
- ... Specify other information depending on the choice above
  including the type of alternative hypothesis (one or two sided).
  For a two sample test, choose Pooled 'on' if the population variances are assumed to be equal, 'off' if they differ.

The calculator produces a test statistic value (t) and a p-value, as well as other information depending on the chosen test.

Use the p-value to determine the outcome of the test at chosen significance level α (reject H_0 if p-value $< \alpha$).

To determine the **critical value** t_d^* for the test statistic t with d degrees of freedom which specifies the rejection region, use the Inverse t-distribution function of the calculator. (Note that this function only calculates upper tails, so will need to change the sign of t^* for a left tailed rejection region.)

The rejection region may be one or two sided depending on H_1. Find t_d^* so that

Two sided	$P(t_d > t_d^*) = \frac{\alpha}{2}$
One sided (right)	$P(t_d > t_d^*) = \alpha$
One sided (left)	$P(t_d < t_d^*) = \alpha$

(Always sketch the rejection region)

To find t_d^* use the methods in Section 9.2.2, but choose F2 t and F3 Invt (instead of F1 NORM and F3 InvN). You will need to specify the number of degrees of freedom for the test.

Example. Have sample information $\bar{x} = 30.6$, $s = 9.5$ and $n = 20$.

Test the hypotheses

$$H_0 : \mu = 25.6 \qquad H_1 : \mu > 25.6$$

A one-sided test.

- F3 TEST
- F2 t (Small sample)
- F1 1-S (One sample)
- $\oplus$ Data; F2 Var
- $\oplus$ μ; F3 $> \mu 0$ (1 sided)

- $\bigoplus \mu 0$; [2] [5] [.] [6] [EXE]
- $\bigoplus \bar{x}$; [3] [0] [.] [6] [EXE]
- $\bigoplus sx$; [9] [.] [5] [EXE]
- $\bigoplus n$; [2] [0] [EXE]
- $\bigoplus$ Execute; [EXE]

Result: Test Statistic, $t = 2.354$, p-value $= 0.01475$.

The degrees of freedom of the test are $n - 1 = 19$

Critical value t^*_{19} for a one-sided $\alpha = 5\%$ test: find t^*_{19} so that $P(t_{19} > t^*_{19}) = 0.05$:

- [F5] DIST
- [F2] t
- [F3] Invt
  (Inverse Normal)
- $\bigoplus$ Area; [0] [.] [0] [5] [EXE]
- $\bigoplus$ df; [1] [9] [EXE]
- $\bigoplus$ Execute; [EXE]

Answer is $t^*_{19} = 1.729$.

The rejection region is $T > 1.729$.

14 Analysis of Variance – ANOVA

Example. Driver reaction times for 3 age groups.

Age Groups		
65+	45-64	25-44
X_1	X_2	X_3
35	18	14
20	10	3
32	21	16
27	14	10
	13	11
		12

Observe differences in their reaction times $\bar{X}_1$, $\bar{X}_2$, $\bar{X}_3$. Are these differences due to different ages (treatment effects) or are they just ordinary variation within drivers?

Enter the data into Lists 1 and 2 in your calculator (STAT mode).

List 2 contains the observations, and List 1 contains the group number of each observation.

e.g. Group 1 contains the observations 35, 20, 32, 27, Group 2 contains 18, 10, 21, 14, 13, and Group 3 contains 14, 3, 16, 10, 11, 12.

	List 1	List 2	List 3	List 4
SUB				
1	1	35		
2	1	20		
3	1	32		
4	1	27		
5	2	18		
6	2	10		
7	2	21		
8	2	14		
9	2	13		
10	3	14		
11	3	3		
12	3	16		
13	3	10		
14	3	11		
15	3	12		

Once the data have been entered:

- F3 TEST
- F5 ANOV
- $\bigoplus$ How many; F1 1
- $\bigoplus$ Factor A; F1 List; 1 EXE
- $\bigoplus$ Dependnt; F1 List; 2 EXE
- $\bigoplus$ Execute; EXE

Gives the first two rows of the ANOVA table, the F statistic value (15.032), and the p-value (0.00053).

	df	SS	MS	F	P
A	2	761.13	380.56	15.032	5.3E-4
ERR	12	303.80	25.316		

Use the p-value to determine the outcome of the test at chosen significance level α (reject H_0 if p-value $< \alpha$).

To determine the **critical value** $F^*_{d1,d2}$ for the test statistic F with $(d1,d2)$ degrees of freedom which specifies the rejection region, use the Inverse F-distribution function of the calculator.

For ANOVA the rejection region is always one-sided. So find $F^*_{d1,d2}$ so that $P(F > F^*_{d1,d2}) = \alpha$.

In the example if $\alpha = 0.01$ then we need $F^*_{(2,12)}$ so that $P(F > F^*_{(2,12)}) = 0.01$:

- F5 DIST
- F4 F (for F distribution)
- F3 InvF
- $\bigoplus$ Data: F2 Var EXE
- $\bigoplus$ Area: 0 . 0 1 EXE (α)
- $\bigoplus$ n:df: 2 EXE; (Numerator df)
- $\bigoplus$ d:df: 1 2 EXE; (Denominator df)
- $\bigoplus$ Execute; EXE

Gives $F^* = 6.9266$: Rejection region is $F > 6.9266$.

15 Chi-Squared Test

Example. Married adults with at least one child reported their family type (intact or other) and whether they experienced violent acts of spouse abuse in the past year.

	Intact	Other
None	743	170
Abuse	36	11

Test whether family type and experience of abuse are associated:

- F3 TEST
- F3 CHI
- F2 2WAY (Use 2-Way because we have crosstabulated 2 variables)
- ⊕ Observed; F2 ▷MAT
- ⊕ Mat A; F3 DIM
- ⊕ m; 2 EXE
- ⊕ n; 2 EXE EXE
- 7 4 3 EXE
- 1 7 0 EXE
- 3 6 EXE
- 1 1 EXE
- EXIT EXIT
- ⊕ Execute; EXE

Result: Test statistic, $\chi^2 = 0.669$, p-value $= 0.41349$.

To view the table of expected values, when looking at the output of the test go

- F6 ▷MAT
- ⊕ Mat B; EXE

You can then check whether the Chi-squared test is valid: it requires 80% of cells to have expected values ≥ 5. Go EXIT EXIT to get back to the test output screen.

To determine the **critical value** χ^{2*}_d for the test statistic X^2 with d degrees of freedom which specifies the rejection region, use the Inverse χ^2-distribution function of the calculator.

For the Chi-squared test the rejection region is always one-sided. So find χ^{2*}_d so that $P(X^2 > \chi^{2*}_d) = \alpha$.

In the example we have $(r-1)(c-1) = (2-1)(2-1) = 1$ degree of freedom. If $\alpha = 0.01$ then we need χ^{2*}_1 so that $P(X^2 > \chi^{2*}_1) = 0.01$:

- F5 DIST
- F3 CHI (for χ^2 distribution)

- F3 InvC
- ⊕ Data: F2 Var EXE
- ⊕ Area: 0 . 0 1 EXE (α)
- ⊕ df: 1 EXE;
- ⊕ Execute; EXE

Gives $\chi^{2*}_1 = 6.6349$: Rejection region is $X^2 > 6.6349$.

16 The Sign Test

Example. Lengths of a certain bone in the skeletons of 9 recently discovered human ancestors were:

$$2.75,\ 2.14,\ 3.23,\ 2.07,\ 2.49,\ 2.28,\ 3.16,\ 2.93,\ 2.20$$

In Australopithicus, the median length of this bone is 2.2 cm.

Test the hypotheses

$$H_0 : m = 2.2 \qquad H_1 : m > 2.2$$

A one-sided test.

To values > 2.2 assign a $+$ (overs)
To values < 2.2 assign a $-$ (unders)
To values $= 2.2$ assign a 0.

Gives: $+-+-++++0$

Drop any 0's and reduce sample size accordingly. Gives $n' = 8$, $2-$ and $6+$.

If H_0 is true then

$$X \sim \text{Binomial}(8, 0.5)$$

Cumulative probability: compute $P(X \geq 6)$:

$$P(X \geq 6) = 1 - P(X \leq 5)$$

- F5 DIST
- F5 BINM
- F2 Bcd
  ('c' for cumulative)
- ⊕ Data; F2 Var
- ⊕ x; 5 EXE
- ⊕ Numtrial; 8 EXE
- ⊕ p; 0 . 5 EXE
- ⊕ Execute; EXE

Answer is $P(X \leq 5) = 0.8555$, so

$$p\text{-value} = 1 - 0.8555 = 0.1445$$

Use the p-value to determine the outcome of the test at chosen significance level α (reject H_0 if p-value $< \alpha$).